Segmentation of Skin Lesions from Digital Images using an Optimized Approach: Genetic Algorithm

Hina Sood¹, Manshi Shukla²

¹M. Tech, Research Scholar, Department of Computer Science and Engineering, RIMT-IET, Mandi Gobindgarh, Fatehgarh Sahib, Punjab, India.
²Assistant Professor, Department of Computer Science and Engineering, RIMT-IET, Mandi Gobindgarh, Fatehgarh Sahib, Punjab, India.

Abstract- The incidence rates of skin cancer diseases Melanoma have been on a rapid increase and is leading to higher death rates. Early detection is of utmost important which increases the survival rates if detected at Stage I. Segmentation of skin lesion is one of the biggest task to do along considering under and over segmentation. An algorithm for detection of skin lesion from digital images using Genetic Algorithm is proposed to get merging of over segmentation to get solved with it. It is one of the most popular techniques to get infinite probabilities solved with genetic algorithm which provides very close results in a short time. The lesion segmentation has been compared to the results of other algorithms. The proposed segmentation algorithm has higher segmentation Sensitivity, Specificity and Accuracy compared to all other segmentation algorithms.

Keywords- Melanoma, Segmentation, Genetic Algorithm, Lesion, Dermoscopy.

I. INTRODUCTION
Melanoma, a deadly skin cancer diseases is a malignant tumor of melanocytes which is causing a large number of deaths and it has been also found that its incidence rates have been on rapid increase [1]. It was estimated that in the United States 1 person among every 49 has the risk of suffering from melanoma [2]. This most deadly form of skin cancer mostly occurs among an age group of 15-30 [3]. However it is been emerging as a leading skin cancers among white-skinned people who come in contact with the ultraviolet radiations [4]. It was also made into notice that melanoma occurs on the lower extremities in a maximum number of cases that have been diagnosed yet [5]. Recent trends have stated that melanoma can be less dangerous if detected at an early stage i.e. if detected at Stage I, the survival rate of the effected person increases to 96% [6].

Due to the increase in the incidence rate of melanoma, researchers are more concerned about proposing such automated systems that diagnose skin lesions accurately and more effectively [7]. Also it has been found that in order to detect melanoma at an early stage screening is beneficial [8]. But the cost of screening melanoma is too high. So, to reduce the screening cost many automated algorithms have been proposed to automatically screen melanoma [8].

A digital dermoscope acquires images that contribute to early screening of melanoma [9] and all automated systems use dermoscopic images [10-11]. A dermoscope is a device that is used to capture images of lesions by the dermatologists. It also magnifies the image and acts as a filter [8]. But most of the practicing dermatologists avoid using dermoscopy due to the lack of training [12]. With dermoscopy it becomes difficult to differentiate malignant and benign lesions and in such case a detailed analysis is needed to be done [13].

Recent studies show that standard digital camera is used to analyze images in automated screening algorithms for melanoma [14]. The standard approach followed in analyzing dermoscopic images automatically includes 1) image segmentation, 2) feature extraction, 3) lesion classification [15]. Segmentation is one of the most important steps in accurately determining a skin lesion. The subsequent steps depend highly on how precisely segmentation has been done. So, segmentation needs to be done with utmost care and is also a difficult task to do as in some lesion images the lesions have an approximately similar intensity value as that of the skin and also the image includes dark hairs covering the lesion. Some examples of such images have been shown in figure 1.

Figure 1: Difficulties in detecting lesions: (a) multiple colored lesions, (b) presence of hair, (c) smooth transition between lesion and skin, (d) specular reflections

Segmentation is needed to be done accurately so that the lesion border is estimated and this segmented result is used further in melanoma detecting algorithms. Many algorithms have been proposed to segment and detect skin lesions automatically among which majority of them are applicable
to dermoscopy images. Some of the automated segmentation algorithms i.e. Independent Histogram Pursuit (IHP) [16], LOG Edge Detector [17], Texture Distinctiveness (TDLS) [18], Iterative Stochastic Region Merging [19] have been summarized under [20]. Many segmentation algorithms use colour information i.e. intensity value of pixels to detect the lesion. The colour information is either used in a single channel or across three colour channels [8]. The various segmentation methods include adaptive thresholding (AT), fuzzy-based split-and-merge algorithm (FBSM) [21] [22], gradient vector flow (GVF) [23], expectation-maximization level set (EM-LS) [24] etc. An another approach to detect skin lesions is to find out the best thresholds with the help of multi-level adaptive thresholding.

In this paper we propose an algorithm based on an optimization technique i.e. Genetic Algorithm (GA) to detect skin lesion in digital images. This algorithm is referred as the Genetic Algorithm based lesion segmentation algorithm. Genetic Algorithm is one of the most popular techniques to get infinite probabilities to get very close results in short time. As we have seen that segmentation of skin lesion is one of the biggest task to do along considering under and over segmentation. One of the biggest hurdles is to merge segmented objects after over segmentation due to its large number of probabilities. So, here we are proposing a unique algorithm to get merging of over segmentation to get solved with Genetic Algorithm. Section II, describes the proposed work. Experimental results are described in Section III. Section IV gives the conclusion of the work done and the future scope.

II PROPOSED WORK

Segmentation of skin lesion is one of the biggest task to do along considering under and over segmentation. One of the biggest hurdles is to merge segmented objects after over segmentation due to its large number of probabilities [8]. So, here we are proposing a unique algorithm to get better results with large number of probabilities in less time.

The first step of the proposed work involves the process of extracting the R, G and B components of the digital image. Through this characterization we represent the respective components as to which shows the lesion more clearly. In the second step Genetic Algorithm incorporates the best thresholds using the fitness function and generates a GA segmented result. Binarization of R, G and B components and Morphological operations and filtration is defined in third step. Figure 2 represents the various steps of the proposed work.

A. Extracting R, G and B Components

The proposed algorithm includes extracting the R, G and B components as its initial step which expresses the three components separately and shows the dissimilarities between the three colours so that we can examine that which component clearly views the lesion. As this initial step firstly the image is loaded in the MATLAB and its R, G and B components are extracted and are shown separately. To describe this step in detail Figure 2 includes some examples that illustrate the process of extraction and also show the components separately.
smoothen the image. Moreover to improve the visual quality and to increase the global contrast of the image Histogram Equalization is used. This adjusts the contrast and makes the image easier to analyze. Then the shadows in the image are removed using Colour Normalization by preserving the colour values.

**B. Applying Genetic Algorithm**

The Genetic Algorithm consists of two main steps. First is Initialization of GA Parameters that initializes the parameters like the number of chromosomes and the number of generations to be operated. The second step Genetic Algorithm is applied to get best thresholds using fitness function. Figure 3 illustrates the overall process of Genetic Algorithm.

Initializing GA Parameters

The first phase of application of Genetic Algorithm is initializing the various parameters that are needed for the computation. The parameters include (a) the number of chromosomes and (b) the number of generations. If we assume the number of chromosomes as ‘C’ and the number of generations as ‘n’ then the algorithm will run for ‘c’ number of chromosomes and will operate on ‘n’ number of generations. More the number of generations more will be the accuracy of the lesion detected. The initialized chromosomes are termed as ‘parent chromosomes’. After the parameters have been initialized the various phases of genetic algorithm are executed to get the best chromosome having best fitness value and generating the GA segmented result. The various steps are as follows:

1) **Mutation**: In the mutation process child chromosomes are generated from the parent chromosomes. It operates at bit level of the image. Generating child chromosomes include changing one if the values among each chromosome with a defined step size accordingly. This results in an equal number of parent and child chromosomes. The parent and child chromosomes share certain characteristics.

2) **Crossover**: Crossover process is used to prepare a child solution with the help of more than one parent. Through this we can vary programming of one chromosome from the other. Crossover process includes crossing over between the best selected chromosomes. In this any one of the value of any of the chromosomes crosses over with the value of some other chromosome.

Applying GA to best threshold using Fitness Function

The second phase of Genetic Algorithm includes the following steps which compute the fitness value and generates best threshold as a final output.

3) **Fitness Value Evaluation**: A fitness function is a function that gives us a fitness value with which we can select the best one for later steps in the algorithm. The maximum value of fitness function denotes the best chromosome. The fitness value evaluation includes evaluating fitness value for both the parent as well as child chromosomes by substituting the value of define parameters into the fitness function. This fitness value is referred to in future in the selection process to extract the best chromosomes.

The fitness function is defined using the weight sum method where the fitness value is generated keeping in concern all the evaluation parameters. The weight sum method initializes equal weight for each parameter used in fitness function which equally increases the value for each i.e. Sensitivity, Specificity and Accuracy. The fitness function so designed for this particular algorithm is given as follows:

\begin{align*}
\text{IR}_1 &= \begin{cases} 
1 & \text{if } (I, j) < \text{thr} \\
0 & \text{otherwise}
\end{cases} \\
\text{IR}_2 &= \begin{cases} 
1 & \text{if } (I, j) = 1 \\
0 & \text{otherwise}
\end{cases} \\
\text{IG}_1 &= \begin{cases} 
1 & \text{if } (I, j) < \text{thg} \\
0 & \text{otherwise}
\end{cases} \\
\text{IG}_2 &= \begin{cases} 
1 & \text{if } (I, j) = 1 \\
0 & \text{otherwise}
\end{cases}
\end{align*}

Figure 4: Flow chart describing Genetic Algorithm.
IB₁ = \{ IB (i, j) < thb \} \text{ then, } \quad IB₂ (i, j) = 1 \\
\text{otherwise, } \quad IB₂ (i, j) = 0 \quad (3)

where \( i \) and \( j \) are the dimensions for image and IR, IG and IB are the indexed component of the image.

Fitness Value = \( (SN(\text{IR}₁|\text{IG}₁|\text{IB}₁)) \times W₁ + (\text{SP}(\text{IR}₁|\text{IG}₁|\text{IB}₁)) \times W₂ + (\text{Acc}(\text{IR}₁|\text{IG}₁|\text{IB}₁)) \times W₃ \) \quad (4)

where \( SN(), \text{ SP()}, \text{ Acc}() \) are the function to evaluate Sensitivity, Specificity and Accuracy and \( W₁, W₂ \) and \( W₃ \) are the respective weights.

4) Selection: The selection process includes selecting the best fitness value that in turn gives best chromosome. The chromosome with a highest fitness value is to be selected. The optimization algorithm deals with the maximum (i.e. maxima) and minimum (i.e. minima) values that generate the best results. Here we are proposing an algorithm that deals with maxima.

5) Elimination: In the elimination phase the chromosomes with least fitness value are replaced with the one having maximum fitness value. This process is repeated for all chromosomes and the chromosomes with least fitness value are eliminated. With this only the best chromosomes are left for further processing and others are eliminated.

After this elimination, the algorithm checks if it has completed its processing for ‘\( n \)’ number of generations. If not, the whole process repeats itself till all generations have been operated.

If the generations are over, the chromosome having best fitness value is chosen and the process is terminated. This generates a best threshold value for the image.

The Genetic Algorithm is summarized as follows in the form of pseudo code.

\[
\text{for } i = 1 \text{ to } C \\
\quad C \_Pᵢ = \text{R} \left( SS \right) \\
\text{END for} \\
\text{for } j = 1 \text{ to } j = n \\
\quad \text{for } i = 1 \text{ to } i = C \\
\quad \text{if } C \_Pᵢ + I(V < 0 \& C \_Pᵢ + I(V < 255 \text{ then } C \_Cᵢ = C \_Pᵢ + I(V) \\
\quad \text{if } C \_Pᵢ + I(V < 0 \text{ then } C \_Cᵢ = C \_Pᵢ \\
\quad \text{if } C \_Pᵢ + I(V > 255 \text{ then } C \_Cᵢ = C \_Pᵢ \\
\text{END for} \\
\text{END for} \\
\text{for } i = 1 \text{ to } i = C \\
\quad C \_Cfᵢ = \text{ff} \left( C \_Cᵢ \right) \\
\text{END for} \\
\text{for } i = 1 \text{ to } i = C \\
\quad \text{if } \left( C \_Cfᵢ > C \_Pfᵢ \right) \\
\quad \quad \left( C \_Pfᵢ > C \_Cfᵢ \right) \\
\quad \quad \left( C \_Pfᵢ > C \_Cfᵢ \right) \\
\text{END if}
\]

END for

Where,
\( R \) is the random function to generate thresholds with SS search limit. \( SS \in [0, 255] \)
\( n \) is the number of generations
\( C \) is the number of chromosomes
\( I(V) \) is the function to mutate chromosomes with \( V \in [-10, 10] \)
\( C \_Cfᵢ \) is the fitness value for its chromosomes
\( \text{ff} \) is the fitness function

C. Binarization of RGB and Morphological Operations

Binarization of RGB using Optimized Thresholds

After applying GA to get best thresholds using fitness function, the binarization of RGB is done using best thresholds. The best thresholds generated by the Genetic Algorithm is that value that gives the best output for the lesion. This value is then converted in binary form in this stage. Thus the optimized thresholds help in detecting the lesion more accurately.

Morphological Operations

After binarization, morphological operations are applied to analyze and process the lesion structure. The basic morphological operators used are dilation, erosion, closing and opening. Here for the proposed algorithm we are using the morphological erosion operation. This operator after application shrinks the objects in image for more correctly detecting the lesion. After applying Morphological Dilation operation, Morphological Filtering is applied to the image for enhancing the image to improve the perceptibility of the image.

C. Summary of the Genetic Algorithm

1) Extract the R, G and B components of the image loaded in the MATLAB.
2) Initialize the Genetic Algorithm by defining the parameters such as the ‘\( C \)’ number of chromosomes to be operated with ‘\( n \)’ number of generations.
3) Computing the best thresholds for the given image.
4) Applying GA to best thresholds using fitness function so that the best value computed by the fitness function is selected for further processing and producing a GA segmented image.
5) Selecting the chromosomes with maximum fitness value and eliminating the others by replacing them by the selected maximum fitness value chromosomes.
6) Checking if all the generation have been operated. If not repeating all the above steps for ‘\( n \)’ number of generations.
7) Storing the optimized threshold.
8) Converting the R, G and B components to their binary forms using optimized thresholds.
9) Predicting outline to the detected lesion.
10) Apply a morphological dilation operation to the resulting image.
11) Give the segmented image with accurately detected and outlined lesion as final image.
III EXPERIMENTAL RESULTS

The proposed work compares the Detected skin lesion using Genetic Algorithm with other segmentation algorithms. The two steps involved in the process are to generate a GA segmented result and the second step is to compute the final resulting image using the best thresholds generated by the genetic algorithm. The Genetic Algorithm is implemented in MATLAB with a Intel (R) Core (TM) i5 CPU (M 480 @ 2.67 GHz 2.66 GHz, 4 GB RAM). To segment a skin lesion in a 1640 * 1043 image, the algorithm computes an average 100% Sensitivity, 99.93% Specificity and 99.94% Accuracy and has an average run time of 45.8893 seconds.

The images used for this algorithm have been taken from the Dermquest database [25]. The images used are corrected using MSIM algorithm [26].

A. GA Segmentation

After the image has been loaded into the MATLAB, and its R, G and B components have been extracted, Genetic Algorithm is applied to the image to get a GA segmented result. This GA segmented result is further used to drive the best thresholds and detect the skin lesion accurately. The GA segmented result of some of the images have been shown as follows:

<table>
<thead>
<tr>
<th>Original Image</th>
<th>GA Segmented</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Original Image" /></td>
<td><img src="image2" alt="GA Segmented" /></td>
</tr>
<tr>
<td><img src="image3" alt="Original Image" /></td>
<td><img src="image4" alt="GA Segmented" /></td>
</tr>
<tr>
<td><img src="image5" alt="Original Image" /></td>
<td><img src="image6" alt="GA Segmented" /></td>
</tr>
</tbody>
</table>

The above table shows the results of the algorithms including the proposed algorithm i.e. Genetic Algorithm (GA).

B. Segmentation Sensitivity, Specificity and accuracy evaluation

The objective of this algorithm is to increase the Sensitivity, Specificity and Accuracy. The metrics that are been used to compare the detected skin lesion to the ground truth are Sensitivity, Specificity and Accuracy. The formulas to compute these three parameters are given in Eqn. 5, 6 and 7.

\[
\text{Sensitivity} = \frac{TP}{TP + FN} \quad (5)
\]

\[
\text{Specificity} = \frac{TN}{TN + FP} \quad (6)
\]

\[
\text{Accuracy} = \frac{TP + TN}{TP + FN + TN + FP} \quad (7)
\]

Here

TP= True Positive i.e. the number of true positive pixels
TN= True Negative i.e. the number of true negative pixels
FN= False Negative i.e. the number of false negative pixels
FP= False Positive i.e. the number of false positive pixels

<table>
<thead>
<tr>
<th>Segmentation Algorithm</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>L-SRM [27]</td>
<td>89.4%</td>
<td>92.7%</td>
<td>92.3%</td>
</tr>
<tr>
<td>Otsu-R [28]</td>
<td>87.3%</td>
<td>85.4%</td>
<td>84.9%</td>
</tr>
<tr>
<td>Otsu-RGB [29]</td>
<td>93.6%</td>
<td>80.3%</td>
<td>80.2%</td>
</tr>
<tr>
<td>Otsu-PCA [30]</td>
<td>79.6%</td>
<td>99.6%</td>
<td>98.1%</td>
</tr>
<tr>
<td>TDLS [8]</td>
<td>91.2%</td>
<td>99.0%</td>
<td>98.3%</td>
</tr>
<tr>
<td>GA</td>
<td>100%</td>
<td>99.93%</td>
<td>99.94%</td>
</tr>
</tbody>
</table>

The above figure shows the graph that compares various segmentation algorithms on the basis of results of Sensitivity, Specificity and Accuracy. Genetic Algorithm shows the highest results.
IV CONCLUSION

A unique algorithm to get merging of over segmentation is proposed. The biggest hurdle was to merge segmented objects after over segmentation due to a large number of probabilities. Also segmentation of skin lesion is one of the biggest task to do along considering under and over segmentation. The proposed algorithm have generated highest result for all the three parameters and has computed very close results in a short time. The future work is worth conducting an analysis of other optimization techniques to produce better results.

ACKNOWLEDGMENT

The author would like to thank the RIMT Institutes, Mandi Gobindgarh, Fatehgarh Sahib, Punjab, India. Author would also wish to thank editors and reviewers for their valuable suggestions and constructive comments that help in bringing out the useful information and improve the content of paper.
REFERENCES


[27] Jeffrey Glaister, Alexander Wong and David A. Clausi “Segmentation of skin lesions from digital images using joint statistical texture distinctiveness”0018- 9294 (c) 2013 IEEE Transactions on Biomedical Engineering.

